

## README file for

### **‘Parental responses to information about school quality: Evidence from linked survey and administrative data’**

Ellen Greaves, Iftikhar Hussain, Birgitta Rabe, and Imran Rasul

## Overview

In this README file we describe the content of the replication package for our paper ‘Parental responses to information about school quality: Evidence from linked survey and administrative data’.

The replication package contains code to replicate our results, the publicly available data used in our project, and full access details about the secure data used in our project.

The code in this replication package constructs the analysis files from five data sources (UK Household Longitudinal Study, Millennium Cohort Study, National Pupil Database, Ofsted management information, School performance tables) using Stata. One master Stata do file and two additional do files (one for the test score analysis using the National Pupil Database and one for the school responses using the Millennium Cohort Study) run all the code to generate the data for 7 original figures and 7 original tables in the paper, and in addition 7 original appendix figures and 9 original appendix tables. The replicator should expect the code to run for about 11.25 hours.

## Data Availability Statements

The main part of the paper combines three different data sets:

- 1) School-level data on school characteristics
- 2) Understanding Society (UKHLS data), secure version
- 3) National Pupil Database (NPD) data linked to Understanding Society

The school-level data (1) are included as part of the replication package. The Understanding Society data (secure version) (2) and linked NPD data (3) can only be used through the Secure Data Service at the UK Data Archive. This is available to registered, non-commercial users who must complete specialist training and be based in the UK when accessing the data. Use of the data requires approval from the data owner or their nominee. Users must apply for access via a Secure Access application and access the data via a secure virtual private network in a safe environment approved by the UK Data Service. Users will be able to upload code and school-level data into the Secure Lab environment.

Two further data sets are used for the analysis (for Tables 7 and Appendix Table 9):

- 1) National Pupil Database data, linked to school characteristics. This can be accessed through the ONS Secure Research Service, under the condition that the researcher is an ONS Approved Researcher and can access the Secure Research Service through

an approved location in the UK. This location can be a ONS Safe Room or Safe Pod, or an office approved as 'Secure' or 'Safe' by the ONS. Access to the National Pupil Database is approved by the Department for Education. It can take some months to negotiate data use agreements and gain access to the data.

- 2) Millennium Cohort Study data, linked to school characteristics. This can be accessed through the Secure Data Service at the UK Data Archive under the same conditions as described above for UKHLS.

#### Statement about Rights

☒ I certify that the author(s) of the manuscript have legitimate access to and permission to use the data used in this manuscript.

☒ I certify that the author(s) of the manuscript have documented permission to redistribute/publish the data contained within this replication package. Appropriate permissions are documented in the [LICENSE.txt](#) file.

#### Summary of Availability

☐ All data **are** publicly available.

☒ Some data **cannot be made** publicly available.

☐ **No data can be made** publicly available.

#### Dataset list

The datafiles included in our replication package are listed below, noting the source

Data file	Source	Notes	Provided
ofsted_ofsted	Ofsted		Yes
SchoolXs	DfE		Yes
school_wider	DfE		Yes
School_Performance_Tables	DfE		Yes
school_panel	CLOSER		Yes
Inspections_data_1	Ofsted		Yes
Edubase_alldata_20170126	DfE		Yes

#### Details on each Data Source

>Data dictionary: variable and value labels

#### SchoolXs.dta

laestab; school ID

fte\_pupils; school size

pct\_pupils\_fsm\_eligible; % pupils eligible for FSM  
pct\_pupils\_eal; % pupils with English as their additional language  
acadyr; academic year  
pct\_5ACEM; % pupils 5 A-C GCSE grades including English & Maths  
pct\_5AC; % pupils 5 A-C GCSE grades  
pct\_5AG; % pupils 5 A-G GCSE grades  
tot\_aps; Average total point score  
pct\_progrE; % pupils making expected progress in English between KS 2 and KS 4  
pct\_progrM; % pupils making expected progress in Maths between KS 2 and KS 4  
pct\_EngBacc; % pupils achieving the English Baccalaureate  
hassixthform; School has a sixth form  
christian; Christian school  
otherrelig; Other religious school  
boys; Boys school  
girls; Girls school  
mixedgender; Mixed gender school  
lea; Local education authority  
sch\_acad; Academy school  
sch\_commun; Community school  
sch\_volun; Voluntary aided school  
sch\_found; Foundation school  
sch\_spec; Special school

**ofsted\_ofsted.dta**

laestab; school ID  
acadyr; academic year  
inspection\_date; inspection date  
inspgr; Ofsted inspection rating

gap; number of years since previous inspection  
status; change to previous inspection rating  
inspgr\_lag; previous Ofsted inspection rating

### **school\_panel.dta**

laestab\_2005; LAESTAB in LEASIS year 2005  
laestab\_2006; LAESTAB in LEASIS year 2006  
laestab\_2007; LAESTAB in LEASIS year 2007  
laestab\_2008; LAESTAB in LEASIS year 2008  
laestab\_2009; LAESTAB in LEASIS year 2009  
laestab\_2010; LAESTAB in LEASIS year 2010  
laestab\_2011; LAESTAB in LEASIS year 2011  
laestab\_2012; LAESTAB in LEASIS year 2012  
laestab\_2013; LAESTAB in LEASIS year 2013

### **school\_wider.dta**

laestab; school ID  
acadyr; academic year  
inspection\_date; inspection date  
inspgr; Ofsted inspection rating  
inspgr\_lag; previous Ofsted inspection rating  
fte\_pupils; school size  
pct\_pupils\_fsm\_eligible; % pupils eligible for FSM  
pct\_5ACEM; % pupils 5 A-C GCSE grades including English & Maths  
pct\_5AC; % pupils 5 A-C GCSE grades  
pct\_5AG; % pupils 5 A-G GCSE grades  
tot\_aps; Average total point score  
pct\_progrE; % pupils making expected progress in English between KS 2 and KS 4

pct\_progrM; % pupils making expected progress in Maths between KS 2 and KS 4

pct\_EngBacc; % pupils achieving the English Baccalaureate

hassixthform; School has a sixth form

christian; Christian school

otherrelig; Other religious school

boys; Boys school

girls; Girls school

mixedgender; Mixed gender school

lea; Local education authority

sch\_acad; Academy school

sch\_commun; Community school

sch\_volun; Voluntary aided school

sch\_found; Foundation school

### **School\_Performance\_Tables.dta**

dfesno: school ID (dfesno)

urn: school Unique Reference Number

year: academic year

school\_type: Community, Academy, Voluntary Aided etc.

school\_religion: religious ethos

scl\_admissions\_policy: comprehensive, selective etc.

scl\_gender\_type: single sex school type

totpups: Number of pupils on roll (all ages)

ttaps: total average point score

p5GCSE\_A\_G: Percentage of pupils achieving 5+ A\*-G or equivalents

ttapscp: Total average (capped) point score per pupil

std\_ttapscp: school-level standardised Total average (capped) point score per pupil

pfsm: % pupils eligible for FSM

pEAL: Percentage of Key Stage 4 pupils with English as an additional language (EAL)

Pexpected\_progress\_KS4: Percentage of pupils at the end of Key Stage 4 achieving expected level of progr

lag\_ttapscp: 1 year lagged total average point score

lag\_std\_ttapscp: 1 year lagged standardised total average point score

lag\_p fsm: 1 year lagged % pupils eligible for FSM

lag\_pEAL: 1 year lagged % EAL

lag\_p5GCSE\_A\_G: % pupils 5 A-C GCSE grades

lag\_Pexpected\_progress\_KS4: 1 year lagged % pupils making expected progress

lag\_totpups: 1 year lagged total pupils

### **Inspections\_data\_1.dta**

urn: school Unique Reference Number

insp\_year: academic year of inspection

inspgr: inspection grade

inspection\_date: date of inspection

primary\_secondary: primary or secondary school

inspection\_start\_date: start date of inspection

inspection\_end\_date: end date of inspection

### **Edubase\_alldata\_20170126.dta**

urn: school Unique Reference Number

LinkURN: link school Unique Reference Number

LinkName: link school name

LinkType: child or parent school

LinkEstablishedDate: date link established

## Computational requirements

### Software Requirements

Stata (code was last run with version 15.1)

- `ssc install outreg2` (as of 10/05/2017)
- `ssc install estout` (as of 30/06/2018)
- `ssc install coefplot` (as of 29/03/2018)
- `ssc install catplot` (as of 15/10/2021)
- `ssc install mmerge` (as of 16/05/2017)

### Memory and Runtime Requirements

The code was last run on a **Intel(R) Core(TM) i5-3570 CPU @ 3.40GHz, 32GB RAM PC, Windows 10 Enterprise** with running time **11 hours and 16 minutes**.

### Description of programs/code

The RMASTER do file runs all code to produce all output aside from Table 7 and Appendix Table 9. Output files are called appropriate names (T1.csv, F4.png) and should be easy to correlate with the manuscript. Note that Appendix Tables are outputted as N-1 (for example in the code TA4 corresponds to TA5 in the paper) as the final manuscript has an additional appendix table 2. All the figures (for example sample sizes) used in the text and notes to Figures are saved to a log file.

Within the MASTER do file:

- Data set-up:
  - The file `$prog\RSchool_level_data.do` cleans the school characteristics and Ofsted ratings.
  - The file `$prog\RSample_selection_children.do` cleans the UKHLS data and creates the main analysis data sets. (This file also creates Appendix Table 2.)
- Analysis on main data sources:
  - The file `$prog\Rdescriptives.do` produces descriptive evidence and figures used in the main paper. (Figures 1, 2, 4, A2, A3; Tables 1, 2, 6, A4, A6.)
  - The file `$prog\Raux_analysis.do` produces auxiliary analysis. (Tables 3, A5.)
  - The file `$prog\Rbootstrap_main.do` produces the main analysis. (Figure 5; Tables 4, 5, A6.)
  - The file `$prog\Rbootstrap_heterogeneity.do` produces the heterogeneity analysis for the main effects. (Figures A6A, A6B, A7.)

- The file \$prog\Rbootstrap\_other.do replicates the main analysis for additional outcomes. (Figure F7.)
- The file \$prog\Rrobustness\_checks.do tests the robustness of the main analysis to alternative specifications and samples. (Figures A4, A5; Tables A8, A11.)
- Analysis on additional data sources:
  - The file \$prog\R\_MCS\_analysis.do uses an alternative dataset (the Millennium Cohort Study) to study teacher/school responses to the inspection outcomes. (Table A9.)
  - The file \$prog\RTest\_scores\_TABLE\_7.do uses the National Pupil Database to study the overall effect on the inspection outcome on pupils' test scores. (Table 7.)

### Instructions to Replicators

- If required, become an ONS Approved Researcher and UK Data Service Accredited Researcher.
- Apply to use the Secure License UKHLS and National Pupil Database for the required purpose.
- Import replication package into secure environment
- Edit \$prog/RMASTER.do to adjust the default path (global prog).
- Run \$prog/RMASTER.do to run all steps in sequence.

### Details

- \$prog/RMASTER.do: will create all output directories, install needed ado packages and produce all original results in the paper.
- The programs were last run top to bottom on 25<sup>th</sup> November, 2021.

### List of tables and programs

The provided code reproduces:

- ☒ All numbers provided in text in the paper
- ☒ All tables and figures in the paper
- ☐ Selected tables and figures in the paper, as explained and justified below.



Figure/Table #	Program	Line No.	Output file	Note
Figure 1	Rdescriptives.do	733, 746, 763, 783	F1A.png, F1B.png, F1C.png, F1D.png	Requires confidential data
Figure 2	Rdescriptives.do	826, 832, 849	F2A.png, F2B.png, F2C.png	Requires confidential data
Figure 3	n/a	n/a	n/a	Research design diagram
Figure 4	Rdescriptives.do	345	F4.png	Requires confidential data
Figure 5	Rbootstrap_main.do	173	F5.png	Requires confidential data
Figure 6	Rdescriptives.do	1007	F6.png	Requires confidential data
Figure 7	Rbootstrap_other.do	125, 135	F7A.png, F7B.png	Requires confidential data
Figure 8	n/a	n/a	n/a	Source: Varkey Foundation
Table 1	Rdescriptives.do	362	T1.csv	Requires confidential data
Table 2	Rdescriptives.do	120	T2.csv	Requires confidential data
Table 3	Raux_analysis.do	85, 97, 112, 127, 143	T3.txt	Replicable

Table 4	Rbootstrap_main.do	104, 114	T4.csv, T4_N.csv	Requires confidential data
Table 5	Rbootstrap_main.do	122	T5.csv	Requires confidential data
Table 6	Rdescriptives.do	484, 540	T6.csv	Requires confidential data
Table 7	RTest_scores_TABLE_7.do	257	pupil_level_results.xlsx	Requires confidential data
Appendix Figure 1	n/a	n/a	n/a	Source: Ofsted
Appendix Figure 2	Rdescriptives.do	43, 49	FA2A.png, FA2B.png	Requires confidential data
Appendix Figure 3	Rdescriptives.do	259	FA3.png	Requires confidential data
Appendix Figure 4	Rrobustness_checks.do	695, 898	FA4B.png, FA4A.png	Requires confidential data
Appendix Figure 5	Rrobustness_checks.do	550, 558, 567, 576	FA5A.png, FA5B.png, FA5C.png, FA5D.png	Requires confidential data
Appendix Figure 6	Rbootstrap_heterogeneity.do	185, 193, 201, 209	FA6AA.png, FA6AB.png, FA6AC.png, FA6AD.png,	Requires confidential data
Appendix Figure 7	Rbootstrap_heterogeneity.do	221, 229, 237, 245, 254, 262	FA6BA.png, FA6BB.png, FA6BC.png, FA6BD.png, FA6BE.png, FA6BF.png	Requires confidential data

Appendix Figure 8	Rbootstrap_heterogeneity.do	275, 283	FA7A.png, FA7B.png	Requires confidential data
Appendix Table 1	n/a	n/a	n/a	Source: Ofsted
Appendix Table 2	n/a	n/a	n/a	Source: Ofsted
Appendix Table 3	RSample_selection_children.do	606	TA2.xlsx	Requires confidential data
Appendix Table 4	Rdescriptives.do	658	TA3.csv	Requires confidential data
Appendix Table 5	Raux_analysis.do	230	TA4.csv	Replicable
Appendix Table 6	Rdescriptives.do	135	TA5.xlsx	Requires confidential data
Appendix Table 7	Rbootstrap_main.do	130	TA6.csv	Requires confidential data
Appendix Table 8	Rrobustness_checks.do	201, 341, 354, 359	TA7_1.csv, TA7_2.csv, TA7_N.csv, TA7_p.csv	Requires confidential data
Appendix Table 9	R_MCS_analysis.do	346	TA8.csv	Requires confidential data
Appendix Table 10	Rrobustness_checks.do	1019, 1052, 1056, 1062, 1066, 1070, 1074	TA9.csv, TA9_R2.txt	Requires confidential data
Appendix Table 11	Rrobustness_checks.do	455, 475	TA10_1.csv, TA10_2.txt	Requires confidential data

## References

### Understanding Society Secure Access dataset (SN 6676)

University of Essex, Institute for Social and Economic Research. (2020). *Understanding Society: Waves 1-10, 2009-2019 and Harmonised BHPS: Waves 1-18, 1991-2009: Secure Access*. [data collection]. 11th Edition. UK Data Service. SN: 6676, <http://doi.org/10.5255/UKDA-SN-6676-11>

### Understanding Society Special License School Codes (SN 7182)

University of Essex, Institute for Social and Economic Research. (2020). *Understanding Society: Waves 1-10, 2009-2019: Special Licence Access, School Codes*. [data collection]. 8th Edition. UK Data Service. SN: 7182, DOI: [10.5255/UKDA-SN-7182-8](http://doi.org/10.5255/UKDA-SN-7182-8)

### Linked NPD (SN 7642)

Department for Education, University of Essex, Institute for Social and Economic Research. (2020). *Understanding Society: Linked Education Administrative Datasets (National Pupil Database), England, 1995-2018: Secure Access*. [data collection]. 3rd Edition. UK Data Service. SN: 7642, <http://doi.org/10.5255/UKDA-SN-7642-3>

### Millennium Cohort Study (SN 8227)

University College London, UCL, Institute of Education, Centre for Longitudinal Studies (2017). *Millennium Cohort Study: Ofsted and Teacher Survey Linked Data, 2008: Secure Access*. UK Data Service: SN: 8227. <http://dx.doi.org/10.5255/UKDA-SN-8227-1>

### School Codes (SN 7182)

University of Essex, Institute for Social and Economic Research. (2021). *Understanding Society: Waves 1-10, 2009-2019: Special Licence Access, School Codes*. [data collection]. 8th Edition. UK Data Service. SN: 7182, <http://doi.org/10.5255/UKDA-SN-7182-7>